

## ANNIE - Support #25532

### Archiver restart in 2021

02/17/2021 03:26 PM - Arthur Kreymer

<b>Status:</b>	Work in progress	<b>Start date:</b>	04/02/2021
<b>Priority:</b>	Low	<b>Due date:</b>	07/04/2021
<b>Assignee:</b>	Arthur Kreymer	<b>% Done:</b>	0%
<b>Category:</b>		<b>Estimated time:</b>	10.00 hours
<b>Target version:</b>		<b>Spent time:</b>	46.00 hours
<b>Description</b>			
The DAQ and offline archiver processes need cleanup and restart after the SL7 related issues that arose in late 2020.			
<b>Subtasks:</b>			
Support # 25705: Deal with the 229 /data/archive-2018 files			Assigned
Support # 25708: ingress file cleanup			Assigned

### History

#### #1 - 02/17/2021 03:38 PM - Arthur Kreymer

- Status changed from Assigned to Work in progress

An issue that caused problems in August 2019 was that Enstore changed from ecrc to adler 32 (ecrc -1) in its Layer 4 metadata. The fts-ftp and stored scripts were not updated to handle this. This interfered with preening files on the DAQ system.

The issues that interrupted the archiver at the SL7 upgrade were :

1. Loss of the annieraw login area offline  
The login area was /nashome/a/annieraw which at SL7 became NFS4 mounted. This mount required a kerberos ticket which the annieraw account does not have. The login area was moved to /annie/app/home/annieraw via RITM1066969 on 2021 Jan 06. Similarly for the anniepro account.
2. Loss of kerberized ftp offline  
Kerberized ftp is used to move files into dCache.  
Under SLF6, this is /usr/kerberos/bin/ftp.  
Under SLF7 that path is missing, and does not seem to be available.  
I copied the ftp binary from the SLF6 annie-daq02 to /annie/app/home/annieraw/archiver/kftp on 2021 Feb 11, and adjusted the offline fts-ftp script.
3. Loss of the offline keytab file  
We use files in /opt/annieraw/ to generate a kerberos principal for file copies.  
annieraw.keytab - the keytab file  
annieraw.principal - the name of the principal in the keytab  
These were restored from annie-da01 on 2021 Feb 11.

#### #2 - 02/19/2021 09:47 AM - Arthur Kreymer

I have put a link to the DAQ Archiver documents under the annie\_experiment wiki

Phase II Computing and Software > Data Management > DAQ Data Archiving

#### #3 - 02/19/2021 10:24 AM - Arthur Kreymer

## PLAN OUTLINE

Generally, run the scripts by hand before restarting in cron.

## This plan has been moved to

<https://cdcv.sfnal.gov/redmine/projects/anniesoft/wiki/RESTART-2021>

## OFFLINE Archiver

- Clear space in persistent, presently 97% full 2021 02 24

- Verify that all persistent/raw files are on tape, using the check-raw script (kreymer)
- Ran fts-ftp to re-archive three inconsistent files. (kreymer) 2021 02 22
- Removed appropriate files from persistent (kreymer) 2021 02 24
- see <https://cdcv.s.fnal.gov/redmine/projects/anniesoft/wiki/CHECK-RAW>
- **Check status of the stored script, which allows files to be preened.** 2021 02 19 (kreymer)
- \* **MOOT - will modify preen to use the Enstore Complete File Listing directly**
  - 02/19 kreymer
  - found 1260 files with empty ftscrc values, all Aug 26,27 2017
  - files are listed in /annie/data/users/annieraw/maint/ftscrcshort-20210219.txt
  - created fixcrc script to update the missing values,
    - did a dry-run not actually updating
    - takes 1 or 2 seconds per file
  - 02/20 kreymer
    - Ran fixcrc for real , then rsyncstored
    - Corrected two file crc values to match Enstore's new adler32 values, reran rsyncstored

## ONLINE archiver (kreymer)

- See why the 1766 files in /data/archive have not been preened
  - 1260 of these are likely due to missing ftscrc values, see notes above
  - Most are preened, 504 remain in /data/archive-2018
    - 493 have mismatched checksums
    - 1 is doing, 10 need to be moved to do
- **Run rsyncstored to prepare for preening** 2021 02 19
- \* **MOOT - will modify preen to use the Enstore Complete File Listing directly**
  - 02/19 kreymer
  - created new version of rsyncstored, to clean up ssh messages
  - ran this. Up to date now
  - 02/29 reran after running fixcrc and stored offline
- **Run preen to clear /data/archive** 2021 02 24 kreymer
  - 1262 files preened
- **Deploy new preen using Complete File Listing** 2021 03 01 (kreymer)
- **Run archdriver 2021 03 08** (kreymer)
- Enable crontab (kreymer)
- Retire do/doing/done, move files from archive to copied (kreymer)

## DAQ ( brichard, lpickard )

- reactivate DAQ processes which produce RAWData and V6Data files

## CLEANUP ( all )

- Remove historical files
  - /annie/data/users/annieraw: ftscrc, ftsum, stored
  - /data/archiver : do doing done stored
- deal with the files moved in 2020/2021 with temporary scripts (brichard)

### #4 - 02/20/2021 03:27 PM - Arthur Kreymer

- % Done changed from 10 to 20

I identified issues with the offline scripts, largely due to a change in the Enstore metadata in August 2019. Formerly an incorrect adler32 checksum was placed into the Layer 4 metadata, The checksum was seeded with 0 rather than 1. The seed was corrected. But Annie's fts-ftp and staged scripts were not adjusted. Older files have the same old checksum, newer files the new checksum.

You get the new checksum by running the ecrc -1 command.

We are now in a position to run preen on the online system, except that there are also hundreds of files with mismatched md5sum checksum.

I would prefer to understand and fix this before removing anything /pnfs/annie/persistent. These errors date back to 2017. See /data/logs/preen.log

### #5 - 02/22/2021 11:11 AM - Arthur Kreymer

I created a Data Operations page at [https://cdcv.s.fnal.gov/redmine/projects/anniesoft/wiki/RAW\\_DATA\\_OPERATIONS](https://cdcv.s.fnal.gov/redmine/projects/anniesoft/wiki/RAW_DATA_OPERATIONS)

I will finish rerunning check-raw as needed on all persistent areas today,

and update that wiki page.

There are 493 files which are not the same online and offline, as noted by preen.

I will add a page for PREEN support today, to summarize the stats of these 493.  
I plan to pre-stage these files to dCache as needed, as many are not on persistent.  
Then an expert can evaluate why they are different.

#### #6 - 02/22/2021 06:18 PM - Arthur Kreymer

I created an offline archiver/preenstage script which classifies the files listed as MISMATCH files in preen.

Some claim to have good checksums, need to look into this.  
Some claim to not be in the tape directory. Need to see if these are in Persistent.  
One is on the fly, rewritten today but is not yet actually on tape, probably waiting for a drive.  
I will check again tomorrow.

#### #7 - 02/24/2021 02:11 PM - Arthur Kreymer

Ben confirmed at the software meeting today that all the /pnfs/annie/raw files are old and not of current interest, and that they can be removed.

As agreed, I have removed all files from /pnfs/annie/persistent/raw .

See details under <https://cdcvns.fnal.gov/redmine/projects/anniesoft/wiki/CHECK-RAW>

#### #8 - 02/24/2021 02:11 PM - Arthur Kreymer

- % Done changed from 20 to 50

#### #9 - 02/24/2021 02:55 PM - Arthur Kreymer

As agreed at today's software meeting,  
I have set aside the 1766 un-preen data files in /data/archive for future study.  
The space used there is 800 GB.

```
cd /data

latest() { find ${1} -type f -print0 | xargs -0 stat --format '%Y :%y %n' \
| sort -nr | cut -d: -f2- | head -3 ; }

$ latest archive
2018-06-21 14:36:07.105919440 -0500 archive/RAWDDataR922S10p0.root
2017-10-03 22:51:06.845868651 -0500 archive/V6DataR897S1p0T6Oct__2.root
2017-10-03 22:32:36.471785187 -0500 archive/V6DataR898S0p0T7Oct__3.root

mv archive archive-2018
mkdir archive

ls -ld archive archive-2018
drwxrwxr-x 2 annie annie 6 Feb 24 14:50 archive
drwxrwxr-x 2 annie annie 262144 Jul 18 2018 archive-2018
```

#### #10 - 02/24/2021 03:09 PM - Arthur Kreymer

- % Done changed from 50 to 60

The Fermilab storage people report that a file had errors being written to tape.  
Is ~~/pnfs/annie/raw/raw/1387/RAWDDataR1387S0p11~~  
~~rw-r--r--~~ 1 annieraw annie 798958310 Nov 20 12:20 /pnfs/annie/raw/raw/1387/RAWDDataR1387S0p11

I do not find this anywhere on annie-daq01, so we cannot recover it.  
The files p0 through p10 for that run seem to be OK.

Ben - shall I simply remove this file from the tape archive ?  
/pnfs/annie/raw/raw/1387/RAWDDataR1387S0p11

#### #11 - 02/24/2021 08:29 PM - Arthur Kreymer

- % Done changed from 60 to 70

I have preened the 1262 files from /data/archive-2018 which are archived on tape.

I identified them from the preentest.log file and moved them back to /data/archive.

I ran a preview of bin/preen - looked OK, appropriate file counts.

Ran the actual preen

```
set nohup ; ${HOME}/bin/preen +7 do &
```

This took about 35 minutes to run, see /data/logs/preen.log

/data/archive-2018 now has the remaining 504 files, 228 GB.

These are the files that have different online and offline checksums.

I will look into these later, at lower priority.

For now, I think we are in good shape to resume normal archiving.

Manually at first, then via cron.

Ben plans to work on producing /data/archive/\*.root files next week.

#### #12 - 02/26/2021 02:43 PM - Arthur Kreymer

There are 11 data files that were never written to the archives.

This should be enough to test the scripts.

I have moved these 11 files back into /data/archive.

I have archived one of these, RAWDataR335S6p58.root and am waiting for it go to tape.

I am also restructuring preen to use the daily Enstore Complete File Listing for comparison, rather than our rather kludged list of stored files with their locally calculated md5sums.

The CFL can be read from the web.

#### #13 - 03/01/2021 02:09 PM - Arthur Kreymer

I did a bit more cleaning up of the archiver scripts.

I preened the one file archived last week.

I archived the remaining 10 files from /data/archive-2018 that did not show checksum errors.

These were vintage Aug 2017.

I removed use of the local md5sum and ecrc copies of file checksums.

These were redundant, as dCache and Enstore will tell you the file checksum, and it is quick and easy to do a checksum on the local file.

I changed the fts-ftp internal integrity check from md5sum to ecrc -1, for consistency with what dCache/Enstore use.

The DAQ archftp script no longer calculates or pushes checksums to offline.

The DAQ system no longer needs to run rsyncstored, and offline does not need to run the stored script.

We now get the file Enstore checksums from the Enstore Complete File Listing, via the web.

That file is only updated daily, but that is fine for preening.

I will keep whittling away at the /data/archive-2018 backlog of 493 files that had issues.

Meanwhile, I think we are ready to accept new current files in /data/archiver.

I have not restored the crontabs.

I would rather run the scripts by hand until I stop making changes.

#### #14 - 03/03/2021 12:07 PM - Arthur Kreymer

- % Done changed from 70 to 80

On Oct 4, 2017, 264 files had zero md5sum content written to /annie/data/users/annieraw/stored with zero content.

This tripped up the preen processes which needed valid md5 sums there.

The files are all on tape with valid sizes and adler32 sums.

I have validated manually that the tape file md5sums match those in the ftssum files written during the initial archive.

So the files are on tape, with valid content.

The preen script no longer uses the stored area.

It should be fine to preen the files on the DAQ system.

I moved the 264 files from /data/archive-2018 to /data/archive for preening.

This leaves the 229 files in /data/archive-2018 which have mismatched md5 sums.

But due to the age of these files,  
the Enstore checksums are based on 'ecrc', not 'ecrc -1' used by preenn

I have run a special preenz script that uses the older ecrc values.

#### #15 - 03/04/2021 03:58 PM - Arthur Kreymer

Ben and Arthur met via Zoom late Thursday morning 03/04  
to discuss the immediate next steps for restarting archiving.

Ben created ~annie/bin/README to summarize the process there.

New files are no longer in root format.  
The archiver and preen scripts should be OK with this.  
Arthur will adjust the offline scripts as necessary.

We will leave the 229 files having uncertain checksums as-is in /data/archive-2018  
These accumulated at various times from June 2016 through August 2017.  
That 133 GBytes of space is a small fraction of the 9 TB /data partition.

Arthur will continue to tune and run the scripts manually,  
probably starting up cron jobs next week.

Ben, Arthur, Leon et.al. will meet in a week or two to review progress and make further plans.

#### #16 - 03/04/2021 04:57 PM - Arthur Kreymer

I updated annidata/archiver/fts-ftp to no longer require a .root file extension.  
I think we are ready to start archiving new data files from /data/archive

I think the old do, doing, done infrastructure is too complicated, and probably not needed.  
Perhaps the archive script should instead move files from /data/archive to /data/copied  
( /data/archived sounds too much like /data/archive )  
I will give this some more thought.  
But we can proceed as-is for now.

#### #17 - 03/05/2021 09:54 AM - Arthur Kreymer

Email from Ben around 05:55 CST :

files are now in the correct directory to start archiving.

Arthur replies :

I have preened the remaining six .root files with  
\${HOME}/bin/preen do

I have archived 1, then 10 files, all looks good.  
I will proceed with 100, then all files.

For example :

```
DAQ
```

```
${HOME}/bin/archdo 10  
less /data/logs/archiver/archdo.log
```

```
${HOME}/bin/archiver  
less /data/logs/archiver/archiver.log
```

```
OFFLINE
```

```
${HOME}/archiver/fts-ftp  
less ${HOME}/archiver/fts-ftp.log
```

#### #18 - 03/05/2021 10:44 AM - Arthur Kreymer

The 1, 10 and 100 files archives have been done.  
The rate writing to dCache is about 400 files per hour.

It will likely take a day to archive the remaining 4000 files

About 5 hours to move to the dropbox, the 10 to write to Enstore.

**#19 - 03/05/2021 10:58 AM - Arthur Kreymer**

- Due date changed from 03/01/2021 to 03/15/2021

**#20 - 03/06/2021 11:57 AM - Arthur Kreymer**

- % Done changed from 80 to 90

I took the liberty of updating the offline fts-ftp script to verify files copied to Enstore by comparing their Layer 2 dCache checksums and lengths, rather than doing an active re-read of the files.

This roughly doubles the speed of archiving. I do not recall a case where the Layer 2 checksums have not been reliable in Fermilab dCache.

The archive was started via  
set nohup ; \${HOME}/archiver/fts-ftp &

This will take several hours to complete

Here is an extract from fts-ftp.log

```
fts-ftp 20210306
LIMIT 9999
Sat Mar 6 10:17:04 CST 2021 on anniegpvm02.fnal.gov
3829/3829 FILES NEEDED
Sat Mar 6 10:17:27 CST 2021 RAWDataR2453S0p112
TIME 35
Sat Mar 6 10:18:05 CST 2021 RAWDataR2453S0p113
Sat Mar 6 10:18:08 CST 2021 RAWDataR2453S0p114
Sat Mar 6 10:18:12 CST 2021 RAWDataR2453S0p115
Sat Mar 6 10:18:17 CST 2021 RAWDataR2453S0p116
Sat Mar 6 10:18:21 CST 2021 RAWDataR2453S0p117
Sat Mar 6 10:18:25 CST 2021 RAWDataR2453S0p118
...
Sat Mar 6 14:53:30 CST 2021 RAWDataR2539S0p8
Sat Mar 6 14:53:38 CST 2021 RAWDataR2539S0p9
Sat Mar 6 14:53:45 CST 2021
```

**#21 - 03/08/2021 09:44 AM - Arthur Kreymer**

I restarted the DAQ archdriver script, which runs the archiver every hour.

```
date ; set nohup ; ${HOME}/bin/archdriver &
```

```
Mon Mar 8 09:39:15 CST 2021
[1] 1600
```

**#22 - 03/08/2021 10:54 AM - Arthur Kreymer**

I cloned the overall plan posted to this issue Feb 19, to a separate Redmine page,

<https://cdcv.s.fnal.gov/redmine/projects/anniesoft/wiki/RESTART-2021>

**#23 - 03/10/2021 11:49 AM - Arthur Kreymer**

Ben - the files in /data/archive are owned by root, so cannot be removed by preen,.

Please fix ownerships.

**#24 - 03/10/2021 12:11 PM - Arthur Kreymer**

The /data/archive directory is owned by annie. So preen can remove the files owned by root by doing rm -f.

I have updated the preen script, and restarted preen. It seems to be working now.

**#25 - 03/10/2021 02:15 PM - Arthur Kreymer**

I will restructure archftp so that we no longer need to run archdriver.

This will remove the need for the 130K + files in /data/archiver/done and stored directories.

The stored directory is already retired. I have copied it to /home/annie/maint/stored.  
After we have met next week, I intend to remove it entirely.

I plan to likewise stash a copy of the too large /data/archiver/done area,  
and remove it when we have met next week.

Once the files are archived and preened, I see no reason to keep these working files around.  
We have the job logs to record the work that was done.

#### **#26 - 03/12/2021 11:57 AM - Arthur Kreymer**

I have enabled the annieraw crontab on anniegpvm02

```
06 03 * * * ${HOME}/archiver/fts-ftp
```

I have removed the need for the archdriver script, and the /data/archiver/do and doing areas.  
preen runs strictly based on local files and the complete file listing  
preen removed preened file flags from the /data/archiver/done  
archiver checks that files are not flagged in /data/archiver/done  
archiver can run daily or as desired from cron, using a pid file to avoid multiple runs

I have stashed a copy of the /data/archiver/done in /home/annie/maint/done  
for reference in understanding the files in /data/archive-2018

I have moved the /data/archiver stash and test areas, vintage 2015, to /home/annie/maint

#### **#27 - 03/14/2021 09:58 PM - Arthur Kreymer**

I have preened manually this morning and afternoon.  
338 files in /data/archive are waiting to age the required 7 days for preening.

I will test and deploy the latest version of archiver when new files show up in /data/archive.  
The latest files there are from March 8.

Once the latest archiver script is tested,  
I will enable the crontab for daily archiver and preen runs.

#### **#28 - 03/17/2021 11:31 PM - Arthur Kreymer**

Ben, Marcus and Arthur discussed the short term plans, as well as FCRSG planning.

Arthur created Redmine Issue to hold FCRSG information for 2021.  
<https://cdcv.sfnal.gov/redmine/issues/25658>

Ben will add Leon to the annepro account access list.  
Ben has started daily movement of files to /data/archive - around 03:00

Arthur will :

1. Have Marcus added to the if-admin-annie access list - RITM1104320
2. Send information about dCache and Enstore quotes - mailed to Ben and added to Issue 25658
3. Start the DAQ archiver in cron after manual tests
  - Delayed due to kerberized ftp failing after today's dCache maintenance - see INC000001113857
  - kftp should be working tomorrow 03/18, we may need a better supported long term solution
4. Clean up the obsolete entries in /data/archiver/done and /home/annie/maint/, and similar offline areas

#### **#29 - 03/18/2021 12:42 PM - Arthur Kreymer**

- Due date changed from 03/15/2021 to 03/19/2021

The dCache developers have fixed their internal firewall for now.  
We are waiting for the changes to be deployed via puppet.  
Then I will restart the archiver on annie-da01.  
I am following INC000001113857 (ftp) and RITM1104773 (deploy fix)

#### **#30 - 03/18/2021 01:29 PM - Arthur Kreymer**

The immediate kftp problem has been resolved.  
I have manually started an archiver run on annie-da01.  
I have enabled the crontab on annie-da01

crontab -l

```
07 20 * * * ${HOME}/bin/preen do
07 04 * * * ${HOME}/bin/archiver
```

preen runs after the daily updates to the Enstore complete file listing.  
archiver runs after the daily movement of files to /data/archive

### #31 - 03/19/2021 12:22 PM - Arthur Kreymer

I adjusted the DAQ crontab to run the archiver just after new files are produced at 03:00

```
07 03 * * * ${HOME}/bin/archiver
07 20 * * * ${HOME}/bin/preen do
```

I adjusted the offline crontab to run after the DAQ pushed new files to the dropbox

```
06 04 * * * ${HOME}/archiver/fts-ftp
```

### #32 - 03/20/2021 11:09 PM - Arthur Kreymer

- % Done changed from 90 to 80

Most files are working fine, but preen found one file different in the archive.

RAWDataR2567S0p7 is 400 bytes longer offline than online.  
Strangely, the differences seem to be at the beginning of the file.  
See details below

The copies in persistent and tape are identical.

I have moved the persistent copy offline to /pnfs/annie/persistent/BAD/raw/RAWDataR2567S0p7  
I have removed the tape copy.

I have removed the /data/archiver/done/RAWDataR2567S0p7 so that a fresh copy will be logged.  
I have re-archived the /data/archive copy.  
I have added an Adler32 checksum value to the /data/archiver/done flag file, formerly empty.

Which version of RAWDataR2567S0p7 is bad - the /data/archive version  
or the copy in /pnfs/annie/persistent/BAD/raw/RAWDataR2567S0p7 ?

Here are some diagnostics :

#### ONLINE

```
$ od /data/archive/RAWDataR2567S0p7 | head
0000000 105437 000010 000000 000000 177400 075534 072163 140545
0000020 062673 016607 133333 055555 155435 106666 111616 035433
0000040 154351 034352 033067 136073 107235 133073 035555 175167
0000060 157367 114633 051465 152577 125251 147265 166051 166775
0000100 126775 034123 177137 074273 134530 114733 155071 114771
0000120 155171 034271 004013 134630 154133 074732 120533 100575
0000140 101606 101606 117770 057461 137540 077774 077745 044067
0000160 045112 074631 102632 021425 157774 167456 074777 106440
0000200 175603 003357 123060 044105 167310 144574 070304 042336
0000220 170365 104617 020567 115306 111341 174375 023211 004115

$ od /data/archive/RAWDataR2567S0p7 | tail
12444141320 120754 006301 077433 035527 152466 161710 153467 005710
12444141340 051331 147443 177051 147447 057231 047467 177250 117131
12444141360 176560 057160 104704 074443 144401 171507 162440 057150
12444141400 155140 036321 037772 166523 175175 163065 000014 000010
12444141420 000000 000000 000000 057555 067546 066562 072141 000004
12444141440 000000 000000 000000 000000 000000 000016 000000 000000
12444141460 000000 057555 074564 062560 064143 061545 064553 063556
12444141500 000001 000000 000000 000000 120400 055301 106745 127001
12444141520 000126
12444141521
```

```
$ wc /data/archive/RAWDataR2567S0p7 | head
 5341309   30630768 1418773329 /data/archive/RAWDataR2567S0p7
```

#### OFFLINE

```
$ od /pnfs/annie/persistent/BAD/raw/RAWDataR2567S0p7 | head
0000000 042507 020124 020057 052110 050124 030457 030456 005015
0000020 071525 071145 040455 062547 072156 020072 067515 064572
0000040 066154 027541 027065 020060 053450 067151 067544 071567
```



```
0000060 047040 020124 030061 030056 020073 033170 024464 040440
0000100 070160 062554 062527 045542 072151 032457 033463 031456
0000120 020066 045450 052110 046115 020054 064554 062553 043440
0000140 061545 067553 020051 064103 067562 062555 033457 027060
0000160 027060 032463 034063 033456 020067 060523 060546 064562
0000200 032457 033463 031456 006466 054012 051455 060543 067156
0000220 071145 020072 062516 071564 060560 065562 071145 005015
```

```
$ od /pnfs/annie/persistent/BAD/raw/RAWDDataR2567S0p7 | tail
12444142140 120754 006301 077433 035527 152466 161710 153467 005710
12444142160 051331 147443 177051 147447 057231 047467 177250 117131
12444142200 176560 057160 104704 074443 144401 171507 162440 057150
12444142220 155140 036321 037772 166523 175175 163065 000014 000010
12444142240 000000 000000 000000 057555 067546 066562 072141 000004
12444142260 000000 000000 000000 000000 000000 000016 000000 000000
12444142300 000000 057555 074564 062560 064143 061545 064553 063556
12444142320 000001 000000 000000 000000 120400 055301 106745 127001
12444142340 000126
12444142341
```

```
$ wc /pnfs/annie/persistent/BAD/raw/RAWDDataR2567S0p7
 5341319   30630798 1418773729 /pnfs/annie/persistent/BAD/raw/RAWDDataR2567S0p7
```

### #33 - 03/20/2021 11:10 PM - Arthur Kreymer

- Due date changed from 03/19/2021 to 04/04/2021

### #34 - 03/22/2021 12:32 PM - Arthur Kreymer

I have reenabled crontab on annie-daq01.

I am drafting a badfile script offline,  
for setting side a corrupt file written to dCache and tape.

We would normally move a copy of bad tape files to /pnfs/annie/BAD/raw.  
The enmv utility which would move a file on tape has stopped working.  
I am in touch with the developers to get a working version.

```
INC000001113991 03/22 enmv error
RITM1106453 03/22 Install encp on Annie interactive systems
```

### #35 - 03/22/2021 01:11 PM - Arthur Kreymer

I updated preen to separately list file size changes when they happen.  
Inspired by the growth in dCache of RAWDataR2567S0p7.

### #36 - 03/23/2021 06:47 PM - Arthur Kreymer

I have seen the Complete File Listing getting updated as late as 20:54.  
So I have adjusted the annie-daq01 crontab entry for preen to 21:07 .

```
cat crontab.dat
07 03 * * * ${HOME}/bin/archiver
07 21 * * * ${HOME}/bin/preen do
```

### #37 - 03/23/2021 06:49 PM - Arthur Kreymer

- % Done changed from 80 to 90

Developers have made available a working version of enmv,  
for setting aside bad files on tape.

I have updated the offline annieraw:archiver/badfile script.

If and when we get another damaged file, I will test this.

### #38 - 03/24/2021 05:58 PM - Arthur Kreymer

The annie-daq01 /home/annie/bin/preen script is updated to send mail  
when there is a mismatched file size or checksum.

This is expected to be rare, monthly or less based on /data/logs/archiver/preen.log

preen sends to the comma separated list /home/annie/bin/MAILTO .  
Presently that list just has one entry, [kreymer@fnal.gov](mailto:kreymer@fnal.gov) .

Please add yourself as desired.

**#39 - 04/02/2021 10:43 AM - Arthur Kreymer**

I have updated the online preen script and the offline ftsftp script to post summaries of recent activity to the web each time they run.

These pages are under <https://anniedata.fnal.gov/data/raw/>

Enjoy !

**#40 - 04/02/2021 11:04 AM - Arthur Kreymer**

- *Start date changed from 02/17/2021 to 04/02/2021*

- *Due date changed from 04/04/2021 to 07/04/2021*

due to changes in a related task: [#25705](#)

**#41 - 04/02/2021 11:14 AM - Arthur Kreymer**

I have created a subtask <https://cdcv.s.fnal.gov/redmine/issues/25705> to track the cleanup of files in /data/archive-2018